

# MASTER'S THESIS

## Het Effect van Retrieval Practice versus Herlezen op de Metacognitive Accuracy van Acht- en Negenjarige Basisschoolleerlingen

Ketelaar, Marill

**Award date:**  
2020

[Link to publication](#)

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### Take down policy

If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 04. May. 2023

**Open Universiteit**  
[www.ou.nl](http://www.ou.nl)



Het Effect van Retrieval Practice versus Herlezen op de Metacognitive Accuracy van Acht- en  
Negenjarige Basisschoolleerlingen

The Effect of Retrieval Practice versus Restudy on the Metacognitive Accuracy of Eight- and Nine  
Year Old Primary School Students

Marill Ketelaar

Master Onderwijswetenschappen  
Open Universiteit

Datum: 05-02-2020

Begeleider: Dr. G. Camp

## Inhoud

Samenvatting .....	3
Summary .....	5
Inleiding .....	6
Methode .....	13
Procedure .....	15
Resultaten .....	17
Conclusie en discussie .....	18
Referenties .....	20

Het Effect van Retrieval Practice versus Herlezen op de Metacognitive Accuracy van Acht- en Negenjarige Basisschoolleerlingen

Marill Ketelaar

**Samenvatting**

Uit verschillende onderzoeken blijkt dat *retrieval practice* (ophalen van kennis uit het geheugen) tot een beter geheugen leidt voor die kennis dan herlezen. Daarnaast blijkt uit onderzoek dat het ook leidt tot een betere *metacognitive accuracy* van jongvolwassenen. Dit omvat het nauwkeurig inschatten van welke leerstof beheerst wordt en welke leerstof nog niet. In onderzoek wordt dit vaak vastgesteld door proefpersonen een *judgement of learning (JOL)* te laten maken. Deze inschatting wordt vervolgens vergeleken met de daadwerkelijke prestatie op de toets. Wanneer de inschatting nauwkeurig is, heeft de leerling een betere basis waarop hij of zij aan zelfregulatie kan doen. De effectiviteit van zelfregulatie is namelijk in grote mate afhankelijk van de nauwkeurigheid van de *metacognitive monitoring*. Leerlingen die een betere inschatting kunnen maken in hoeverre zij de leerstof beheersen, kunnen hun leertijd effectiever kunnen besteden. *Metacognitive accuracy* wordt ook geassocieerd met betere schoolprestaties en een beter geheugen. Onduidelijk was of *retrieval practice* óók een positief effect heeft op de *monitoring accuracy* van kinderen in de basisschoolleeftijd.

Doel van dit onderzoek was om inzicht te verkrijgen in het effect van *retrieval practice* in vergelijking met de strategie herlezen op de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen.

Er participeerden 20 acht-/negenjarigen in het onderzoek. Deze kinderen zitten op een openbare basisschool gelegen in het dorp Zelhem en zijn geselecteerd op basis van hun leeftijd. Er werd gebruik gemaakt van een *within-subjects* design, waarbij de gehanteerde *within-subjects factor* de door participanten gebruikte oefenstrategie (*retrieval practice* vs. herlezen) is. De afhankelijke variabele was de Goodman- en Kruskal Gamma ( $\gamma$ ) correlatie tussen *JOL* en latere prestatie op een *item-by-item* toets. De hypothese was dat *retrieval practice* leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij acht- en negenjarigen.

Het woordenschatniveau van de kinderen werd vastgesteld met behulp van Cito woordenschat 3.0 voor groep 5 (Cito). Daarnaast werd er in het onderzoek gebruik gemaakt van 20 woordparen die gebruikt worden als leerstof en tevens als toetsmateriaal.

Uit de analyse bleek dat *retrieval practice* ( $M = .86, SE = .20$ ) gemiddeld genomen leidde tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen ( $M = .67, SE = .33$ ). Dit

## Het Effect van Retrieval Practice versus Herlezen op de Metacognitive Accuracy

verschil, .18, BCa 95% CI [.06, .31], was significant  $t(18) = 3.00, p = .008$ . Het representeert een middelsterk positief effect,  $d = 0,69$ .

Uit dit onderzoek bleek dat *retrieval practice* leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen. Op basis van deze uitkomsten verkrijgt het gebruik van de leerstrategie *retrieval practice* de voorkeur boven herlezen, omdat het de *metacognitive accuracy* van kinderen bevordert. Omdat *metacognitive accuracy* geassocieerd wordt met betere schoolprestaties en een beter geheugen kan het toepassen van deze aanbeveling van ten goede komen voor het basisonderwijs.

Trefwoorden: *retrieval practice, judgement of learning, metacognitive accuracy*

### Summary

Various studies show that retrieval practice, or retrieving knowledge from your memory, leads to a better memory for that knowledge than re-reading. In addition, research shows that it also leads to better metacognitive accuracy of young adults. This includes accurately assessing which subject material is mastered and which subject material is not yet mastered. In research, this is often determined by having participants make a judgment of learning (JOL). This estimate is then compared with the actual performance on the test. When the estimate is accurate, the student has a better basis on which he or she can self-regulate their study behaviour. This in turn leads to more effective learning. Metacognitive accuracy is therefore associated with better school performance and better memory. It is unclear whether retrieval practice also has a positive effect on the monitoring accuracy of children of primary school age.

The aim of this research was to gain insight into the effect of retrieval practice in comparison with the strategy re-read on the metacognitive accuracy of eight and nine year old primary school pupils.

Twenty eight-/nine-year-olds participated in the study. These children attend a public primary school in the village of Zelhem and were selected based on their age. An within-subjects design was used, whereby the within-subjects factor was the study strategy used by participants (retrieval practice vs. re-reading). The dependent variable was the Goodman and Kruskal Gamma ( $\gamma$ ) correlation between JOL and subsequent performance on an item-by-item test. The hypothesis was that retrieval practice would lead to a more accurate assessment of one's own performance on the final test than re-reading among eight and nine year olds.

The vocabulary level of the children was determined using Cito woordenschat 3.0 for third grade (Cito). In addition, the research used 20 word pairs that are used as learning material and test material.

The analysis showed that retrieval practice ( $M = .86$ ,  $SE = .20$ ) on average lead to more accurate predictions of final test performance than rereading ( $M = .67$ ,  $SE = .33$ ). This difference, .18, BCa 95% CI [.06, .31], was significant  $t(18) = 3.00$ ,  $p = .008$ . It represents a medium strong positive effect,  $d = 0.69$ .

This study showed that *retrieval practice* lead to a more accurate leads to more accurate performance predictions than rereading for eight-/nine-year-olds. Based on the results of this study, the use of retrieval practice is preferred over rereading, because retrieval practice foster *metacognitive accuracy* of children. Because *metacognitive accuracy* is associated with better schoolperformance and better memory, applying this recommendation can be beneficial for primary education.

Keywords: *retrieval practice, judgement of learning, metacognitive accuracy*

## Inleiding

### Probleemschets en doel van het onderzoek

Hoewel er vanuit de cognitieve psychologie veel bekend is over de effectiviteit van bepaalde leerstrategieën (zie Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013), leren we als leerlingen/studenten niet vaak over de manier waarop die strategieën in de praktijk kunnen toepassen. In het huidige onderwijs worden een aantal strategieën gebruikt die niet of nauwelijks effectief blijken, waaronder het herlezen van teksten. Dit terwijl andere strategieën die wél effectief blijven, zoals *retrieval practice*, het ophalen van kennis uit het geheugen, veel minder vaak gebruikt worden (Blasiman, Dunlosky, & Rawson, 2017; Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007; Morehead, Rhodes, & DeLozier, 2016). Uit onderzoek naar de effectiviteit van *retrieval practice* bij kinderen blijkt dat deze strategie een positief effect heeft op het onthouden van woorden door negenjarigen (Karpicke, Blunt, & Smith, 2016). Daarnaast blijkt dat *retrieval practice* achtjarigen helpt bij het aanleren van woordenschat (Goossens, Camp, Verkoeijen, Tabbers, & Zwaan, 2014) en het geheugen van negen- en tienjarigen verbetert op het gebied van topografie (Rohrer, Taylor, & Sholar, 2010). *Retrieval practice* heeft dus effect op het geheugen van kinderen, maar heeft *retrieval practice* ook een effect op hun metacognitie?

Metacognitie kan beschreven worden als ‘kennis over de eigen kennis’. Om effectief het eigen leren te kunnen reguleren is de *metacognitive accuracy*, de nauwkeurigheid van de metacognitie, van groot belang. Meerdere onderzoeken hebben het belang van *metacognitive accuracy* vastgesteld voor de onderwijspraktijk. Zo wordt het geassocieerd met betere schoolprestaties (Everson & Tobias, 1998; Nelson et al., 1994) en een beter geheugen (Thiede, Anderson, & Theriault, 2003; Thiede, Dunlosky, Griffin, & Wiley, 2005).

Gezien deze voordelen is er onderzoek gedaan naar hoe *metacognitive accuracy* kan worden verbeterd (Kelemen, Winningham, & Weaver, 2007; Hacker, Bol, Horgan, & Rakow, 2000; Thiede et al., 2003). Zo wordt in meer specifieke en recentere studies onderzocht in hoeverre de bewezen effectieve oefenstrategie *retrieval practice* effect heeft op metacognitie (Carpenter, 2012; Dunlosky et al., 2013; Karpicke, Lehman, & Aue, 2014; Kornell & Vaughn, 2016; Roediger & Karpicke, 2006b; Roediger, Putnam, & Smith, 2011; Rowland, 2014). Uit onderzoek van Boettcher (2013) blijkt dat *retrieval practice* de *metacognitive monitoring accuracy* van jongvolwassenen verbetert. Daarnaast blijkt uit onderzoek van Little en McDaniel (2015) dat de inschatting van prestatie op de toets van universiteitsstudenten accurater zijn wanneer zij de oefenstrategie *retrieval practice* gebruikten tijdens het leerproces, dan wanneer zij de oefenstrategie herlezen gebruikten.

Deze uitkomsten impliceren dat *retrieval practice* een effectievere oefenstrategie is dan herlezen als het gaat om het verbeteren van de *metacognitive accuracy*. Echter, deze onderzoeken zijn enkel gericht op jongvolwassenen. Het doel van dit onderzoek is om inzicht te verkrijgen in het al dan niet

verbeteren van de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen door gebruik van de oefenstrategie *retrieval practice* in tegenstelling tot herlezen.

### **Retrieval practice**

Jezelf testen met *flash cards*, het doen van een quiz of een bingo zijn verschillende manieren om leerstof te oefenen. Wat deze werkvormen gemeen hebben is dat informatie uit het geheugen opgehaald wordt, een proces dat het langetermijngeheugen verbetert. Deze oefenstrategie wordt ook wel *retrieval practice* genoemd. Deze oefenstrategie blijkt veel effectiever dan veelgebruikte oefenstrategieën zoals herlezen en arceren (bijvoorbeeld, Roediger & Butler, 2011). *Retrieval practice* omvat elke activiteit waarbij men geleerde leerstof ophaalt uit het geheugen nog vóór een daadwerkelijke toets waarmee gemeten wordt in hoeverre de leerstof is beklijfd. Het ophalen van de leerstof kan zowel opzettelijk als incidenteel zijn en de leerstof kan verschillende typen informatie bevatten. Zo kan het gaan om een specifieke gebeurtenis, een procedure of een feit (Fazio & Marsh, 2018).

De voordelen van *retrieval practice* zijn vooral onderzocht voor adolescenten en volwassenen, zowel in het laboratorium als in de klas (Roediger, Agarwal, McDaniel, & McDermott, 2011). Hoewel er minder onderzoek is gedaan naar de voordelen van *retrieval practice* bij jonge kinderen, blijft dit onderzoeksveld wel groeien (Fazio & Marsh, 2018). Tot op heden blijkt uit onderzoek naar de effectiviteit van *retrieval practice* bij kinderen dat deze strategie een positief effect heeft op het onthouden van woorden door negenjarigen (Karpicke et al., 2016). Daarnaast blijkt dat *retrieval practice* het aanleren van woordenschat verbetert bij achtjarigen (Goossens et al., 2014) en het geheugen van negen- en tienjarigen verbeterd op het gebied van topografie (Rohrer et al., 2010).

Hoewel veel onderzoek zich richt op de voordelen van *retrieval practice*, is er minder onderzoek gedaan naar de reden wáárom deze oefenstrategie leidt tot het verbeteren van het langetermijngeheugen. Een mogelijke verklaring hiervoor is dat door de overeenkomst in het format tussen *retrieval practice* (de oefentoets) en de eindtoets de stof gemakkelijker kan worden opgehaald na *retrieval practice* op de eindtoets dan na herbestuderen. Echter, verschillende onderzoeken ontkrachten dit omdat situaties waarbij de oefentoets en eindtoets meer op elkaar leken niet leidde tot betere resultaten op de eindtoets. De beste resultaten op de eindtoets werden behaald door participanten die als oefentoets zo veel mogelijk behandelde leerstof uit het geheugen ophaalden, ook wel *free recall* genoemd (bijvoorbeeld Carpenter & DeLosh, 2006; Kang, McDermott, & Roediger, 2007). Deze uitkomsten suggereren dat het *retrieval practice* effect niet primair verklaard kan worden door de gelijkenis tussen de oefentoets en eindtoets, maar door ophaalprocessen die plaatsvinden tijdens de oefentoets. Een andere mogelijke verklaring is dat het ophalen helpt om onderscheid te maken tussen leerstof die al wel beheerst wordt en leerstof die nog niet goed genoeg beheerst wordt,



waardoor *retrieval practice* invloed heeft op het leren tijdens een volgend leermoment (Roediger & Karpicke, 2006a). Maar zelfs als er geen corrigerende feedback of extra blootstelling aan de leerstof plaatsvindt, blijkt dat het ophalen van informatie op een oefentoets effectiever is dan het herlezen (bijvoorbeeld Carpenter & DeLosh, 2005; Cull, 2000; Kuo & Hirshman, 1996; Roediger & Karpicke, 2006b). Carpenter (2009) zocht naar een verklaring door te onderzoeken wat voor soort verwerking er optreedt tijdens het ophalen, wat waarschijnlijk niet gebeurt tijdens het herlezen. In deze experimenten werden woordparen die ofwel sterk ofwel zwak geassocieerd zijn geleerd door middel van *retrieval practice* of herlezen. Uit de eindtoets bleek dat participanten die gebruik gemaakt hadden van *retrieval practice* de woordparen beter onthouden hadden dan de participanten die de woordparen hadden herlezen. Hoewel sterk geassocieerde woordparen het ophalen aanvankelijk makkelijker maakte, werden zwak geassocieerde woordparen die geleerd werden door middel van *retrieval practice* beter onthouden over een langere tijd. Op basis van de uitkomsten van dit onderzoek veronderstelt Carpenter dat *retrieval practice* het geheugen kan verbeteren door het teweegbrengen van uitgebreide ophaalprocessen. Wanneer men informatie probeert op te halen uit het langetermijngeheugen wordt gerelateerde informatie geactiveerd. Deze geactiveerde informatie wordt samen met de opgehaalde doelinformatie gecodeerd, wat zorgt voor een 'spoor' dat meerdere paden mogelijk maakt om latere toegang tot die informatie te vergemakkelijken (Carpenter, 2009). Pyc en Rawson (2010) veronderstellen een vergelijkbare verklaring, namelijk dat *retrieval practice* de codering van effectieve *mediators* vergemakkelijkt. Dit houdt in dat *retrieval practice* ervoor zorgt dat diep verwerkte informatie verbonden wordt aan aanwijzingen en doelen (Pyc & Rawson, 2010).

### **Metacognitie**

Flavell (1979) gebruikte de term metacognitie voor het eerst en definieerde het als de kennis die iemand heeft over de kwaliteit en nauwkeurigheid van de eigen cognitie. Nelson et al. (1994) stelden een kader op voor metacognitie, waarin zij stellen dat verwerking van leerstof plaatsvindt op twee niveaus: het objectniveau (*metacognitive control*) en het metaniveau (*metacognitive monitoring*). Het objectniveau houdt in dat men gericht is op de uit te voeren taak, terwijl het metaniveau inhoudt dat men gericht is op het monitoren van de cognitieve processen die bij de taak betrokken zijn (Nelson et al., 1994). *Metacognitive monitoring* houdt in dat leerlingen hun eigen leerprestaties bijhouden, zodat zij vooruitgang zien en duidelijk wordt wat het effect is van hun inspanning (Zimmerman, 1990).

### **Zelfregulerend leren**

*Metacognitive monitoring* is een voorwaarde voor zelfregulerend leren (bijvoorbeeld Winne & Hadwin, 1998; Zimmerman, 2008). Kenmerken van zelfregulerend leren zijn het gebruikmaken van

effectieve strategieën om het leerproces te reguleren, het reflecteren op de effectiviteit van het leren en de motivatie om te leren. Ter verduidelijking zal een situatie geschetst worden.

Nick heeft over vijf dagen een toets geschiedenis. Hij moet hiervoor 20 begrippen leren met bijbehorende betekenis. Nick begint met het maken van flitskaarten. Aan de voorkant zet hij het begrip, aan de achterkant de betekenis van het begrip. Hij legt alle flitskaarten met de begrippen naar boven op tafel. Van een aantal begrippen weet hij de betekenis meteen. Deze flitskaarten legt hij apart. Hij gaat door totdat hij de betekenissen van de meeste begrippen een keer goed heeft genoemd. Zes begrippen krijgt hij maar niet in zijn hoofd. Hij maakt hierbij een tekening. De volgende dag doet Nick hetzelfde met de flitskaarten. Hij maakt een stapel van begrippen die hij al goed kent en een stapel van begrippen waarbij hij af en toe nog twijfelt over de betekenis. De laatstgenoemde stapel oefent hij nogmaals. De dag erop doet Nick weer hetzelfde. De stapel met begrippen waarbij hij af en toe twijfelt wordt steeds kleiner. Op de vierde dag dat Nick leert voor de toets, herhaalt hij alleen nog de begrippen waarbij hij af en toe nog twijfelde over de betekenis. Uiteindelijk kent hij alle begrippen. De laatste dag laat hij zich overhoren met behulp van de flitskaarten door een vriend.

Leerlingen die effectief strategieën inzetten om hun eigen leerproces te reguleren, hebben de neiging om hogere leerprestaties te laten zien, hebben meer motivatie om te leren en kunnen beter problemen oplossen en omgaan met tegenslagen (Zimmerman, 1990).

### **Metacognitive accuracy**

De effectiviteit van de zelfregulatie hangt in hoge mate af van de nauwkeurigheid van de *metacognitive monitoring*. Leerlingen die nauwkeurig kunnen inschatten welke leerstof zij beheersen en welke leerstof nog niet, kunnen effectiever zelfregulerend leren doordat zij zich focussen op dat wat nog beter geleerd moet worden om goed te presteren op een aankomende toets. Dit in tegenstelling tot leerlingen die hier niet toe in staat zijn. Zij besteden hun leertijd eerder inefficiënt, doordat zij tijd en energie steken in het leren van leerstof die zij al beheersen. Ook lopen zij het risico om leerstof niet goed genoeg te bestuderen, omdat zij in de veronderstelling zijn dat zij het beheersen. Hierdoor kunnen prestaties op een toets lager uitvallen dan verwacht (Miller & Geraci, 2014).

In verschillende onderzoeken is aangetoond dat een hogere *metacognitive accuracy* kan leiden tot een effectievere regulatie van leergedrag, wat op zijn beurt het leerresultaat kan verbeteren (bijvoorbeeld Thiede, 1999; Thiede & Anderson, 2003; Thiede et al., 2003). Bij veel studies waarin onderzoek gedaan wordt naar het verbeteren van *metacognitive accuracy* moesten participanten voorspellingen doen over hun prestaties (Kelemen et al., 2007).

### Judgement of Learning

Het voorspellen van een prestatie wordt vaak gedaan door middel van *judgements of learning (JOLs)* die ofwel meteen na het aanbieden van de leerstof worden gedaan, ofwel enige tijd na het aanbieden van de leerstof (bijvoorbeeld Arbuckle & Cuddy, 1969; Koriat, 1997, Rhodes & Castel, 2008). Een *JOL* is een voorspelling die gedaan wordt over de waarschijnlijkheid dat recent bestudeerde leerstof later nog herinnerd wordt. In een typisch experiment krijgen participanten leerstof aangeboden, bijvoorbeeld woordparen (apart-bijzonder). Meteen er na of na enige tijd worden participanten gevraagd een *JOL* te maken van de kans dat ze later het synoniem van een woord op kunnen halen na het horen van het eerste woord (apart-?). Na het maken van *JOLs* voor elk item, krijgen participanten een toets over de bestudeerde items. De kwaliteit van de *JOLs* kunnen dan worden beoordeeld op basis van de overeenkomst tussen de inschatting van de prestatie voorafgaand aan de toets (de *JOL*) en de daadwerkelijke prestatie op de toets. Hoe nauwkeuriger de voorspelling de daadwerkelijke leerprestaties weerspiegelen, hoe hoger de nauwkeurigheid van de metacognitieve monitoring (bijvoorbeeld Schraw, 2009). De nauwkeurigheid van de *JOL* kan op twee manieren worden vastgesteld. Allereerst kan de relatieve nauwkeurigheid berekend worden door Goodman- en Kruskal Gamma ( $\gamma$ ) correlaties tussen *JOLs* en latere prestaties op een *item-by-item* toets. Een tweede manier om de nauwkeurigheid van de *JOL* vast te stellen is door de absolute nauwkeurigheid te berekenen door de gemiddelde *JOL* te vergelijken met de gemiddelde prestatie op de toets (Rhodes, 2016).

Meerdere factoren beïnvloeden de nauwkeurigheid van *JOL*. *Delaying JOLs* (het uitstellen van *JOLs*) na de eerste leersessie is de meest effectieve en betrouwbare methode om de nauwkeurigheid van de *JOL* te verbeteren. In een meta-analyse vonden Rhodes en Tauber (2011) dat de gemiddelde gewogen nauwkeurigheid van vertraagde *JOLs*, gemeten op basis van item-correlaties tussen de voorspellingen en de daadwerkelijke geheugenprestaties, .77 was, terwijl deze slechts .44 bedroeg voor directe *JOLs*. Vertraging verbetert niet alleen de nauwkeurigheid van *JOLs* bij volwassenen, maar ook bij basisschoolleerlingen (Schneider et al., 2000; Van Loon et al., 2013). Uit een meta-analyse van Rhodes en Tauber (2011) blijkt dat juist vooral kinderen de grootste voordelen hebben van *the delayed JOL-effect*. Naast *the delayed JOL-effect*, heeft het aantal keren dat de leerstof aangeboden wordt invloed op de nauwkeurigheid van *JOL*. Zo kan de leerstof vaker worden herlezen (Lovelace, 1984; Mazzoni, Cornoldi & Marchitelli, 1990).

Daarnaast is de vormgeving van de toets van invloed. In de meeste onderzoeken wordt er een *cued recall test* afgenomen bij participanten, waarbij bijvoorbeeld het synoniem van een bepaald woord opgehaald moet worden (bijvoorbeeld Dunlosky & Nelson, 1994; Koriat et al., 2006; Nelson & Dunlosky, 1991; Serra, Dunlosky, & Hertzog, 2008). Daarnaast zijn er ook meerkeuzetoetsen (Dunlosky & Nelson, 1997; Thiede & Dunlosky, 1994) en *free recall* toetsen afgenomen. (Dunlosky et al., 2002; Matvey, Dunlosky, & Schwartz, 2006). Uit een meta-analyse blijkt dat de *effect sizes* voor

de gamma correlaties het grootst waren in onderzoeken waarbij een *cued recall* test afgenomen werd (Rhodes & Tauber, 2011).

Tot slot kunnen er oefenstrategieën ingezet worden in de leersessies die van invloed kunnen zijn op de nauwkeurigheid van de *JOL*, bijvoorbeeld dat *retrieval practice* resulteert in een hogere *JOL*-nauwkeurigheid vergeleken met het simpelweg bestuderen van leerstof (Lovelace, 1984; Mazzoni et al., 1990).

### **De effecten van retrieval practice op de metacognitive accuracy**

In verschillende onderzoeken is onderzocht in hoeverre de oefenstrategie *retrieval practice* effect heeft op metacognitie. In het onderzoek van Rawson, O'Neil en Dunlosky (2011) werd onderzocht of *retrieval practice* kan leiden tot het beter onthouden van belangrijke informatie uit patiëntfolders over diabetes. In de eerste leersessie werden de geselecteerde passages uit de patiëntfolders aangeboden ter inzage. In de tweede sessie, een leersessie van 50 minuten, vond de manipulatie plaats. Participanten in de eerste groep moesten zelf leerstof selecteren voor herlezen. In de tweede groep kregen participanten steeds een vraag in beeld, moesten op basis van hun gegeven antwoord een *JOL* doen en kregen de kans om een passage nogmaals te bestuderen. De derde groep kreeg alleen de vragen met bijbehorende relevante passages, maar zonder het doen van een *JOL*. Zij konden zelf kiezen of zij één of meer passages wilden herlezen, of verder wilden gaan naar de volgende vraag. Uitkomst van dit onderzoek is dat de *metacognitive accuracy* toenam na *retrieval practice*. Participanten wisten door *retrieval practice* beter te selecteren welke stof zij nog moesten beter moeten leren voorafgaand aan de daadwerkelijke toets. Hierdoor was de prestatie op de toets beter. Participanten in dit onderzoek waren tussen de 17 en 39 jaar oud (Rawson et al., 2011).

In het onderzoek van Boettcher (2013) werd de *metacognitive accuracy* van eerste generatie studenten (studenten waarvan de ouders geen diploma hebben behaald) en andere generatie studenten (studenten waarvan de ouders wel een diploma hebben gehaald) vergeleken, waarbij ook onderscheid gemaakt werd in het al dan niet gebruiken van de oefenstrategie *retrieval practice*. De participanten in dit onderzoek waren jongvolwassenen tussen de 18 en 26 jaar oud. Uit dit onderzoek bleek dat gebruik van de oefenstrategie *retrieval practice* de *metacognitive monitoring accuracy* voor zowel eerste- als andere generatie studenten verbeterde (Boettcher, 2013).

Het onderzoek van Miller en Geraci (2014) was opgebouwd uit drie experimenten. In het eerste experiment moesten participanten woordparen leren en vervolgens een *JOL* doen. Aansluitend kregen de participanten in de experimentele groep één of twee *retrieval practice* opdrachten, in tegenstelling tot de controlegroep, en moesten daarna nogmaals een voorspelling doen. Hieruit bleek dat *retrieval practice* opdrachten op zichzelf de *metacognitive accuracy* niet verbeterden, maar het falen op deze opdrachten wel. In het tweede experiment werd daarom de moeilijkheidsgraad van de *retrieval*

*practice* opdrachten gemanipuleerd. Uit dit experiment bleek dat het falen op *retrieval practice* opdrachten de *metacognitive accuracy* eerder verbeterde dan wanneer de opdrachten succesvol doorlopen werden. In experiment 3 werd dit nader onderzocht door vier condities te onderscheiden: geen *retrieval practice*, makkelijke *retrieval practice*, *retrieval practice* met een gemiddelde moeilijkheidsgraad en moeilijke *retrieval practice*. Hieruit bleek dat participanten in laatstgenoemde conditie zichzelf vaak onderschatten. Uitkomst van het onderzoek was dat een gemiddelde moeilijkheidsgraad van *retrieval practice* opdrachten de *metacognitive accuracy* positief beïnvloeden. Ook in dit onderzoek waren de participanten jongvolwassenen tussen de 18 en 29 jaar (Miller & Geraci, 2014).

Tot slot hebben Little en McDaniel (2015) experimenten uitgevoerd waarbij participanten teksten moesten lezen. Participanten in de experimentele conditie moesten deelnemen aan een tussentijdse *retrieval practice* toets, in tegenstelling tot de controlegroep die de teksten moest herlezen. Vervolgens moesten zij een *JOL* doen, kregen zij de kans om teksten nogmaals te bestuderen voordat zij de daadwerkelijke toets moesten maken. Uit dit onderzoek bleek dat de inschatting van prestatie op de toets van universiteitsstudenten accurater was wanneer zij de oefenstrategie *retrieval practice* gebruikte hadden tijdens het leerproces, dan wanneer zij de oefenstrategie herlezen gebruikten.

De besproken onderzoeken impliceren dat *retrieval practice* een effectievere oefenstrategie is dan herlezen als het gaat om het verbeteren van de *metacognitive accuracy* van jongvolwassenen. Uit eerder onderzoek is gebleken dat het metacognitief gedrag van kinderen te vergelijken is met het metacognitief gedrag van volwassenen, indien het gebruikte materiaal en de instructies geschikt zijn voor de leeftijd (bijvoorbeeld Metcalfe & Finn, 2012; Krebs & Roebbers, 2012; Schneider, Visé, Lockl, & Nelson, 2000).

### Vraagstellingen en hypothesen

Het doel van dit onderzoek was om inzicht te verkrijgen in het al dan niet verbeteren van de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen door gebruik van de oefenstrategie *retrieval practice* in tegenstelling tot herlezen. De vraag die centraal stond in dit onderzoek was: ‘Wat is het effect van *retrieval practice* in vergelijking met herlezen bij het leren van woordparen op de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen?’ Deze centrale vraag werd onderzocht door middel van experimenteel onderzoek. Gezien de positieve effecten van de oefenstrategie *retrieval practice* op de *metacognitive accuracy* van jongvolwassenen (Rawson et al., 2011; Boettcher, 2013; Miller & Geraci, 2014; Little & McDaniel, 2015), de voordelen van *retrieval practice* bij jonge kinderen vanaf acht jaar (bijvoorbeeld Goossens et al., 2014; Rohrer et al., 2010) en de voordelen van het maken van een vertraagde inschatting voor kinderen vanaf acht jaar (Schneider et al., 2000; Van Loon et al., 2013; Rhodes & Tauber, 2011), werd verwacht dat de

oefenstrategie *retrieval practice* een positieve invloed zou hebben op de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen ten opzichte van herlezen. De volgende hypothese werd daarbij getoetst: '*Retrieval practice* leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen.' Naast deze hoofdhypothese, werd verwacht dat *retrieval practice* zou leiden tot een betere prestatie op de eindtoets dan herlezen (zie Adesope, Trevisan, & Sundararajan, 2017; Dunlosky et al., 2013). De volgende hypothese werd daarbij getoetst: '*Retrieval practice* leidt tot hogere prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen.'

### Methode

#### Ontwerp

In dit onderzoek werd gebruik gemaakt van een experimenteel *within-subjects* design. Kwantitatieve data en gegevens werden door middel van herhaalde metingen verkregen uit één groep participanten. De gehanteerde *within-subjects factor* is de door participanten gebruikte oefenstrategie. De onafhankelijke variabele oefenstrategie bevatte twee condities: *retrieval practice* en herlezen. Alle participanten werden aan beide experimentele condities blootgesteld. Op deze manier konden de prestaties op de ene experimentele conditie, *retrieval practice*, vergeleken worden met de prestaties op de andere experimentele conditie, herlezen (Creswell, 2014). De afhankelijke variabele was de Goodman- en Kruskal Gamma ( $\gamma$ ) correlatie tussen *JOL* en latere prestatie op een *item-by-item* toets. Het voordeel van dit ontwerp is dat de interne validiteit niet beïnvloed wordt door selectie of instrumentatie (Creswell, 2014). Er werden op verschillende manieren zo veel mogelijk factoren uitgesloten die de uitkomsten kunnen beïnvloeden. Allereerst werd er een *pretest* wordt uitgevoerd waarin werd vastgesteld of participanten de synoniemen van de woorden al wisten. Ook werd *counterbalancing* toegepast, wat inhoudt dat de materialen die met *retrieval practice* en herlezen worden geoefend varieerden. Wel konden er specifieke gebeurtenissen plaatsvinden tussen de metingen die mogelijk de uitkomsten van de metingen beïnvloed hebben. Zo kan er tussentijds, buiten de experimentele context, incidenteel extra geoefend zijn met de woordparen. Echter, groepsleerkrachten konden het experiment niet beïnvloeden omdat zij geen voorkennis hadden over het experiment. Tot slot kon de ene experimentele conditie de andere conditie beïnvloeden. Dit werd zo veel mogelijk voorkomen door de instructies over de te gebruiken oefenstrategie zo duidelijk mogelijk te formuleren. Daarnaast werd er zo goed mogelijk op toe gezien dat de instructies opgevolgd werden.

#### Onderzoeksgroep

Er werden 35 leerlingen van een Nederlandse basisschool geselecteerd voor dit onderzoek. Het betreft een openbare basisschool gelegen in het dorp Zelhem. De kinderen die geselecteerd zijn om deel te

nemen zijn 8 tot 9 jaar. De ouders/verzorgers van de kinderen moesten voorafgaand aan het onderzoek een *informed consent* geven voor participatie. De kinderen die mochten deelnemen aan het onderzoek werden vooraf ingelicht dat zij participeren in een experiment. Het woordenschatniveau van de kinderen werd vastgesteld met behulp van Cito woordenschat 3.0. De participanten werden toebedeeld aan één van de *counterbalancing* condities op basis van woordenschatniveau. Deze procedure zorgde ervoor dat eventuele verschillen tussen de twee groepen op de afhankelijke variabele niet werden veroorzaakt door verschillen in woordenschatniveau.

### Materialen

In dit onderzoek werd gebruik gemaakt van Cito woordenschat 3.0 voor groep 5. De toets bestaat uit 40 opgaven die op papier worden aangeboden. Er komen verschillende soorten opgaven voor in de toets. Zo zijn er opgaven met betekenis en betekenisrelaties, maar ook opgaven met en zonder context (Cito, 2018). Vanaf groep 5 bevatten de opgaven alleen tekst. Een voorbeeldopgave is: ‘Welk woord moet op de streep staan? De meester van groep 7 is ziek. Daarom komt er een andere meester \_\_\_\_.’ Er kan gekozen worden uit drie antwoordopties: bijvallen, invallen of opvallen. De kinderen moesten het juiste antwoord omcirkelen.

Daarnaast werd er gebruik gemaakt van de 20 woordparen (bijvoorbeeld: apart-bijzonder) uit het onderzoek van Goossens et al. (2014). Deze woordparen zijn geselecteerd uit een taalboek van groep 7. De moeilijkheidsgraad van de woorden zijn in Goossens et al. (2014) gemeten door gebruik te maken van het meetinstrument ‘*Measure of Lexical Richness*’ (Schrooten & Vermeer, 1994; Vermeer; 2000). Hierin wordt gekeken naar de mediaan van de woordfrequentie in Nederlandse tekstboeken voor basisschoolleerlingen (Schrooten & Vermeer, 1994). Uitkomst hiervan was dat negentien van de 20 woorden in de lijst van Schrooten en Vermeer voorkwamen (1994). De mediaan woordfrequentie van deze 19 woorden was 3 op een schaal van 1 tot 17, wat relatief laag is (Goossens et al., 2014). De woordparen werden digitaal aangeboden als leerstof en als *cues* voor de tussentijdse toetsen en de eindtoets. De woordparen zijn weergegeven in Tabel 1.

Tot slot werd er gebruik gemaakt van een puzzelboekje met rekenkundige opdrachten waaronder sudoku's, doolhoven, kruispuzzels en sommen. Deze was door de onderzoeker een samengesteld.

Tabel 1

*Woordparen*

Woord	Synoniem
apart	bijzonder
baret	mut
beduusd	verrast
chaos	rommel
deponeren	gooien
gift	cadeau
heengaan	doodgaan
heimelijk	stiekem
kris	mes
kwiek	blij
meedelen	vertellen
perplex	verbaasd
pronkstuk	mooi
signaal	teken
stug	stijf
vaal	grauw
vermoeden	idee
verprutsen	verknoeien
wenen	huilen
weezinwekkend	lelijk

### Procedure

De kinderen werden getest in het klaslokaal. Het experiment bestond uit twee sessies. De eerste sessie omvatte de leersessie, gevolgd door een toetsessie één week later. De leerstof werd steeds gepresenteerd op het computerscherm.

Allereerst werd er een schriftelijke pre-test afgenomen waarin de onderzoeker de kinderen vraagt naar de betekenis van de 20 woorden. Na de pre-test startte de leersessie bestaande uit zeven fasen. De eerste en de tweede leerfasen namen samen ongeveer 20 minuten in beslag. De derde tot en met de zevende leerfase waarin de woordparen geoefend werden, namen samen ongeveer 15 minuten in beslag.

In de eerste leerfase vertelde de onderzoeker de kinderen dat zij eerst zouden gaan luisteren naar een woordlijst en vervolgens de betekenissen van de woorden te horen zouden krijgen. De onderzoeker las de 20 woorden op zonder enige toelichting en zonder de synoniemen te geven.

In de tweede leerfase las de onderzoeker de woorden opnieuw op voor de kinderen, terwijl de woordlijst gepresenteerd werd op het computerscherm. De kinderen konden de woorden meelezen terwijl de onderzoeker de woorden toelichtte door het synoniem te noemen en een omschrijving van het woord te geven in één zin.



In de derde leerfase bestudeerden de kinderen alle 20 woordparen en oefenden deze. Ieder woordpaar werd 8 seconde op het scherm getoond, in willekeurige volgorde. De kinderen moesten de woordparen hardop opnoemen.

In de vierde leerfase leerden de kinderen alle 20 woordparen opnieuw. Na deze leerfase was er een korte pauze van twee minuten waarin de kinderen werkten in een puzzelboekje. Dit puzzelboekje bevatte rekenkundige opdrachten, zodat de kinderen afgeleid werden van de taalkundige opdracht met de woordparen.

In de vijfde leerfase kregen de kinderen een blad met daarop de helft van de woordparen. Deze moesten zij herlezen. Vervolgens werden zij getoetst op de andere helft woorden. Welke 10 woorden dit waren is afhankelijk van de *counterbalancing* conditie. De woorden in de herleesconditie werden bestudeerd samen met de bijbehorende synoniemen, net zoals in de vorige leerfase. De woorden in de *retrieval practice* conditie werden weergegeven op een blad, bijvoorbeeld: apart - \_\_\_\_\_. De kinderen werden gevraagd om het synoniem op te schrijven op de lijn. Hiervoor kregen zij 80 seconden de tijd. Na deze fase was er weer een korte pauze van twee minuten waarin de kinderen in het puzzelboekje werkten.

In de zesde leerfase moesten de kinderen allemaal de 20 woorden opnieuw bestuderen op dezelfde manier als in de tweede leerfase. Hierna volgde weer een pauze van twee minuten waarin de kinderen werkten in hun puzzelboekje.

De zevende leerfase was gelijk aan de vijfde leerfase. De kinderen moesten dezelfde 10 woorden herlezen en moesten van de andere 10 woorden het synoniem op de lijn schrijven.

Eén week later vond de toetsessie plaats. Voorafgaand aan de toets gaven de kinderen per woord aan in hoeverre zij dachten het synoniem nog te weten op een schaal van 0 (ik weet het zeker niet) tot 10 (ik weet het zeker wel). De kinderen noteerden deze inschatting per woord op een blad. Vervolgens kregen de kinderen een schriftelijke *cued recall test*. Zij kregen een punt als een synoniem fonetisch identiek was aan het geleerde synoniem. Voor elk goed antwoord kregen de kinderen een punt. Een fout antwoord leverde geen punt op. De toetsen werden afgenomen in het klaslokaal. Er was geen tijdslimiet.

### Analyse

De *judgements of learning* werden beoordeeld op basis van de overeenkomst tussen de inschatting van de prestatie voorafgaand aan de toets en de daadwerkelijke prestatie op de toets. De nauwkeurigheid van de gemaakte *judgement of learning* is in dit onderzoek vastgesteld door de *relative accuracy* te berekenen met behulp van Goodman- en Kruskal Gamma ( $\gamma$ ) correlaties tussen *JOLs* en latere prestaties op een *item-by-item* toets. Dit is een niet-parametrische associatie-index die varieert van -1,0 tot 1,0. Een positieve gamma houdt in dat op de toets herinnerde items hoge *JOLs* kregen en items die

niet onthouden werden lagere *JOLs* kregen. Een negatieve gamma houdt in dat de items die later worden onthouden lage *JOLs* kregen en items die niet werden onthouden hogere *JOLs* kregen (Rhodes, 2016).

De statistische toetsen in dit onderzoek werden uitgevoerd met IBM SPSS Statistics 22. Omdat één groep participanten blootgesteld werd aan twee experimentele condities, werd er een gepaarde *t*-toets uitgevoerd (Field, 2013). Voorafgaand aan deze toets werd eerst gecontroleerd of er wordt voldaan aan de assumptie van normaliteit. Indien hieraan voldaan werd, werd de hypothese getoetst door de groepen te vergelijken. Er werd een *p*-waarde van .05 gehanteerd. Dit betekent dat de nulhypothese (*Retrieval practice* leidt niet tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen) verworpen zou worden bij een *alpha* van 5%. Indien de gepaarde *t*-toets significant was, was er een verschil gevonden tussen de twee groepen. Dit zou inhouden dat er verwacht wordt dat het verschil in gamma's tussen de twee groepen veroorzaakt werd door de gebruikte oefenstrategie.

### Resultaten

Van de 35 leerlingen die geselecteerd zijn, hebben twintig leerlingen daadwerkelijk geparticipeerd in het onderzoek. De ouders van de vijftien overige leerlingen gaven geen toestemming voor participatie aan het onderzoek. De volgende hypothese is getoetst: '*Retrieval practice* leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen.' Er is een gepaarde *t*-toets uitgevoerd om het verschil in de inschatting van de eigen prestatie op de toets vast te stellen voor de leerstrategieën *retrieval practice* en herlezen. De nauwkeurigheid van de gemaakte *judgement of learning* door participanten is vastgesteld door de *relative accuracy* te berekenen met behulp van Goodman- en Kruskal Gamma ( $\gamma$ ) correlaties tussen *JOLs* en latere prestaties op een *item-by-item* toets. Deze techniek is eerder gebruikt in onder andere onderzoek van Thiede et al. (2003), Little en McDaniel (2015) en de Bruin et al. (2011). Eén datapunt is niet meegenomen in de analyse, omdat het niet mogelijk was om een gammacorrelatie te berekenen. Dit was het geval omdat de participant alle woordparen goed had.

Uit de analyse bleek dat *retrieval practice* ( $M = .86, SE = .20$ ) gemiddeld genomen leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen ( $M = .67, SE = .33$ ). Dit verschil, .18, BCa 95% CI [.06, .31], was significant  $t(18) = 3.00, p = .008$ . Het representeert een middelsterk positief effect,  $d = 0.69$ .

De gemiddelde score op *retrieval practice* items was 90.5% ( $SE = 2.094$ ) op de eerste tussentijdse toets en 94.2% ( $SE = 1.170$ ) op de tweede tussentijdse toets.

De gemiddelde toetsscore op *retrieval practice* items op de eindtoets was 46.8% ( $SE = 2.382$ ). Voor de herleesitems gold een gemiddelde toetsscore van 47.9% ( $SE = 2.175$ ) op de eindtoets.

Naast deze hoofdhypothese, werd verwacht dat *retrieval practice* zou leiden tot een betere prestatie op de eindtoets dan herlezen. Gemiddeld genomen is de toetsscore op de *retrieval practice* items (46.8%,  $SE = 2,382$ ), lager dan de toetsscore op de herleesitems (47.9%,  $SE = 2,175$ ). Dit verschil, 0.11, BCa 95% CI [-1.12, -1.33], is niet significant  $t(18) = 0.18, p = .86$ . De hypothese werd verworpen.

### Conclusie en discussie

De hoofdvraag in dit onderzoek was wat het effect van *retrieval practice* in vergelijking met herlezen is bij het leren van woordparen op de *metacognitive accuracy* van acht- en negenjarige basisschoolleerlingen. De eerste hypothese die hieruit afgeleid is stelt dat *retrieval practice* leidt tot het accurater inschatten van de eigen prestatie op de eindtoets dan herlezen bij 8- en 9-jarigen. Deze hypothese wordt in dit onderzoek bevestigd. De *judgements of learning* die gemaakt werden voorafgaand aan de eindtoets waren accurater voor de *retrieval practice* items dan voor de herleesitems. Dit onderzoek verbindt eerdere bevindingen op het gebied van *retrieval practice* (bijvoorbeeld Karpicke et al., 2016; Goossens et al., 2014) en *metacognitive accuracy* (bijvoorbeeld Everson & Tobias, 1998; Nelson et al., 1994; Thiede et al., 2003; Thiede et al., 2005). Daarnaast breidt dit onderzoek eerdere bevindingen uit van de voordelen van *retrieval practice* op de *metacognitive accuracy* bij jongvolwassenen (bijvoorbeeld Boettcher, 2013; Little & McDaniel, 2015) tot de voordelen van *retrieval practice* op de *metacognitive accuracy* bij basisschoolleerlingen. Voor zo ver bekend is bij de onderzoekers, is dit het eerste onderzoek dat de voordelen van *retrieval practice* op het maken van een accurate inschatting van de eigen prestatie aantoonst bij basisschoolleerlingen.

De tweede hypothese was dat *retrieval practice* zou leiden tot een betere prestatie op de eindtoets dan herlezen. Deze hypothese werd in dit onderzoek verworpen. Dat is vreemd, gezien het overweldigende bewijs voor het voordeel van *retrieval practice* ten opzichte van herlezen in de literatuur (bijvoorbeeld, Adesope, et al., 2017; Dunlosky et al., 2013), ook bij jonge kinderen vanaf 8 jaar (bijvoorbeeld Goossens et al., 2014; Rohrer et al., 2010).

Eén mogelijke verklaring is dat voorkennis van invloed is geweest op de toetsscore. Om vast te stellen of kinderen al voorkennis hadden van de materialen die zijn gebruikt, is een pre-test afgenomen. Hieruit blijkt dat van de in totaal 400 getoetste woordparen er vijf al bekend waren. Omdat er maar 1.25% bekend was kan geconcludeerd worden dat voorkennis over de woordparen geen invloed heeft gehad op de toetsscore.

Een andere mogelijke verklaring is dat de items niet goed zijn opgehaald tijdens de leerfasen. Uit de scores op de tussentijdse toetsen (90.5% en 94.2%) blijkt echter dat dit niet het geval is. Wat wel opvalt is dat de score op de eindtoets relatief lager is dan de score op de tussentijdse toetsen. Dit is te

verklaren doordat de eindtoets een week na de leerfasen plaatsvond. Hoewel de snelheid en de mate waarin dingen vergeten worden sterk afhankelijk is van verschillende omstandigheden, is de lagere score op de eindtoets te verklaren door het natuurlijke proces van vergeten (bijvoorbeeld Ebbinghaus, 1880).

Daarnaast kan de manier van toetsen tijdens de eindtoets van invloed zijn geweest. Uit onderzoek blijkt dat *short answer tests*, waaronder de in dit onderzoek gebruikte *cued recall test* onder valt, niet goed kunnen uitpakken bij kinderen jonger dan elf jaar (Goossens et al., 2016) wanneer zij niet van feedback werden voorzien (Lipko-Speed, Dunlosky, & Rawson, 2014). Dit omdat het ophalen van de *cue* in het geheugen beheersing en concentratie vergt van het kind, dat kan worden beïnvloed door afleiding. Aangezien de cognitieve controle nog volop in ontwikkeling is bij kinderen, geldt dit vooral nadeel van *cued recall tests* vooral voor hen (Davidson et al., 2006). Recente bevindingen suggereren echter dat *cued recall tests* onder verdeelde aandachtsomstandigheden nog steeds betrouwbare testeffecten opwekken (Mulligan & Picklesimer, 2017), dus is het de vraag of dit de afwezigheid van het verwachtte positieve effect van *retrieval practice* kan verklaren.

Dit onderzoek is van maatschappelijke relevantie omdat het leidt tot aanbevelingen voor de onderwijspraktijk waardoor de kloof tussen de cognitieve psychologie en praktische toepassingen in de klas verkleind kan worden. Op basis van de uitkomsten van dit onderzoek verkrijgt het gebruik van de leerstrategie *retrieval practice* de voorkeur boven herlezen, omdat het de *metacognitive accuracy* van kinderen bevordert. Omdat *metacognitive accuracy* geassocieerd wordt met betere schoolprestaties (Everson & Tobias, 1998; Nelson et al., 1994) en een beter geheugen (Thiede et al., 2003; Thiede et al., 2005) kan het toepassen van deze aanbeveling ten goede komen voor het basisonderwijs.

Voor toekomstig onderzoek is het interessant om de effecten van *retrieval practice* op de *metacognitive accuracy* van basisschoolleerlingen verder te onderzoeken. Hierbij kan gedacht worden aan jongere kinderen (bijvoorbeeld, 4-5 jaar), maar ook aan andersoortige leerstof dan woordparen. Daarnaast kan er onderzocht worden wat het effect is van verschillende manieren van toetsen op de *metacognitive accuracy* van basisschoolleerlingen.

## Referenties

- Adesope, O. O., Trevisan, D. A., Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659-701.  
doi:10.3102/0034654316689306
- Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, 81, 126–131. doi:10.1037/h0027455
- Boettcher, P. (2013). The Effects of Retrieval Practice on Metacognitive Monitoring Accuracy: A Comparison of First- and Other-Generation Students. *The Journal of Undergraduate Research*, 11, Article 4. Retrieved from  
<https://pdfs.semanticscholar.org/2866/9c18f2adc4d960d7e8bfcf64d652de47cc90.pdf>
- Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: comparing intended versus actual study behaviour. *Memory*, 25(6), 784–792.  
doi:10.1080/09658211.2016.1221974
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1563–1569. doi:10.1037/a0017021
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, 21(5), 279–283. doi:10.1177/0963721412452728
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19(5), 619-636. doi:10.1002/acp.1101
- Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished **cue** support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34(2), 268-276. doi:10.3758/bf03193405
- Cito (2018). *Woordenschat 3.0*. Arnhem: CITO.
- Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston: Pearson.
- Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14(3), 215-235.  
doi:10.1002/(sici)1099-0720(200005/06)14:3<215::aid-acp640>3.0.co;2-1
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037–2078.  
doi:10.1016/j.neuropsychologia.2006.02.006

- de Bruin, A.B., Thiede, K.W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, 109(3), 294-310. doi:10.1016/j.jecp.2011.02.005.
- Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, 33(4), 545– 565. doi:10.1006/jmla.1994.1026
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. (2013). What works, what doesn't. *Scientific American Mind*, 24(4), 47–53. doi:10.1155/2012/741713.
- Ebbinghaus, H. (1880). Urmanuskript "Ueber das Gedächtniß". Passau: Passavia Universitätsverlag.
- Everson, H.T., & Tobias, S. (1998). The ability to estimate knowledge and performance in college: A metacognitive analysis. *Instructional Science*, 26(1-2), 65-79. doi:10.1023/A:1003040130125.
- Fazio, L. K., & Marsh, E. J. (2019). Retrieval-Based Learning in Children. *Current Directions in Psychological Science*, 28(2), 111-116. doi:10.1177/0963721418806673
- Field, A. (2013). *Discovering Statistics using IBM SPSS statistics*. Londen: Sage.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist*, 34(10), 906-911. doi:10.1037/0003-066x.34.10.906
- Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., & Zwaan, R. A. (2014). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition*, 3, 177–182. doi:10.1016/j.jarmac.2014.05.003
- Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92(1), 160–170. doi:10.1037/0022-0663.92.1.160
- Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, 19(1), 126-134. doi:10.3758/s13423-011-0181-y
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4-5), 528–558. doi:10.1080/09541440601056620
- Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrievalbased learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology*, 7, Article 350. doi:10.3389/fpsyg.2016.00350
- Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation*, 61, 1-313.

doi:10.1016/b978-0-12-800283-4.00007-1

- Kelemen, W. L., Winningham, R. G., & Weaver, C. A. III., (2007). Repeated testing sessions and scholastic aptitude in college students' metacognitive accuracy. *European Journal of Cognitive Psychology*, 19(4-5), 689-717. doi: 10.1080/09541440701326170.
- Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, 14(2), 219-224. doi:10.3758/bf03194055\_
- Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of learning and motivation*, 65, 183-215. doi:10.1016/bs.plm.2016.03.003
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349-370. doi:10.1037/00963445.126.4.349
- Koriat, A., Ma'ayan, H., Sheffer, L., & Bjork, R. A. (2006). Exploring a mnemonic debiasing account of the underconfidence-with-practice effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3), 595-608. doi:10.1037/0278-7393.32.3.595
- Krebs, S. S., & Roebbers, C. M. (2012). The impact of retrieval processes, age, general achievement level, and test scoring scheme for children's metacognitive monitoring and controlling. *Metacognition and Learning*, 7(2), 75-90. doi:10.1007/s11409-011-9079-3
- Kuo, T., & Hirshman, E. (1996). Investigations of the testing effect. *American Journal of Psychology*, 109(3), 451-464. doi:10.2307/1423016
- Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 756-766. doi:10.1037//0278-7393.10.4.756
- Lipko-Speed, A., Dunlosky, J., & Rawson, K. A. (2014). Does testing with feedback help grade-school children learn key concepts in science? *Journal of Applied Research in Memory and Cognition*, 3(3), 171-176. doi:10.1016/j.jarmac.2014.04.002
- Little, J. L., & McDaniel, M. A. (2015). Metamemory monitoring and control following retrieval practice for tekst. *Memory & Cognition*, 43(1), 85-98. doi:10.3758/s13421-014-0453-7
- Matvey, G., Dunlosky, J., & Schwartz, B. (2006). The effects of categorical relatedness on judgments of learning (JOLs). *Memory*, 14, 253-261. doi:10.1080/09658210500216844
- Mazzoni, G., Cornoldi, C., & Marchitelli, G. (1990). Do memorability ratings affect study-time allocation? *Memory & Cognition*, 18(2), 196-204. doi:10.3758/bf03197095
- Metcalfe, J., & Finn, B. (2012). Hypercorrection of high confidence errors in children. *Learning and Instruction*, 22(4), 253-261. doi:10.1016/j.learninstruc.2011.10.004

- Miller, T. M., & Geraci, L. (2014). Improving metacognitive accuracy: How failing to retrieve practice items reduces overconfidence. *Consciousness and Cognition*, 29, 131-140. doi:10.1016/j.concog.2014.08.008
- Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, 24(2), 257-271. doi:10.1080/09658211.2014.1001992
- Mulligan, N. W., & Picklesimer, M. (2017). Attention and the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(6), 938-950. doi:10.1037/xlm0000227
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science*, 2(4), 267-270. doi:10.1111/j.1467-9280.1991.tb00147.x
- Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, 5(4), 207-213. doi:10.1111/j.1467-9280.1994.tb00502.x
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335. doi:10.1126/science.1191465
- Rawson, K. A., O'Neil, R., & Dunlosky, J. (2011). Accurate monitoring leads to effective control and greater learning of patient education materials. *Journal of Experimental Psychology: Applied*, 17(3), 288-302. doi:10.1037/a0024749.
- Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. In J. Dunlosky & S. K. Tauber (Eds.), *The oxford handbook of metamemory* (pp. 65-80). New York: Oxford University Press.
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, 137, 615-625. doi:10.1037/a0013684
- Rhodes, M. G. & Tauber, S. K. (2011). The influence of Delaying Judgements of Learning on Metacognitive Accuracy: A Meta-Analytic Review. *Psychological Bulletin*, 137(1), 131-148. doi:10.1037/a0021705
- Roediger, H. I., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15, 20-27. doi:10.1016/j.tics.2010.09.003
- Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210. doi:10.1111/j.1745-6916.2006.00012.x
- Roediger, H. L., & Karpicke, J. D. (2006b). Test enhanced learning: Taking memory tests improves long term retention. *Psychological Science*, 17(3), 249-255. doi:10.1111/j.1467-9280.2006.01693.x



- Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of learning and motivation*, 55, 1-36.  
doi:10.1016/B978-0-12-387691-1.00001-6
- Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 233–239.  
doi:10.1037/a0017678
- Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin*, 140(6), 1432-1463. doi:10.1037/a0037559
- Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's monitoring: Evidence from a judgment-of-learning. *Cognitive Development*, 15(2), 115-134.  
doi:10.1016/S0885-2014(00)00024-1
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, 4(1), 33-45. doi:10.1007/s11409-008-9031-3
- Serra, M. J., Dunlosky, J., & Hertzog, C. (2008). Do older adults show less confidence in their monitoring of learning? *Experimental Aging Research*, 34, 379–391.  
doi:10.1080/03610730802271898
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin and Review*, 6(4), 662-667. doi:10.3758/bf03212976
- Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, 28(2), 129-160.  
doi:10.1016/S0361-476X(02)00011-5
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66-73.  
doi:10.1037/0022-0663.95.1.66.
- Thiede, K. W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology*, 86, 290–302. doi:10.1037/0022-0663.86.2.290
- Thiede, K. W., Dunlosky, J., Griffin, T. D., & Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning Memory, and Cognition*, 31(6), 1267-1280. doi:10.1037/0278-7393.31.6.1267
- van Loon, M. H., de Bruin, A. B. H., van Gog, T., & van Merriënboer, J. J. G. (2013). The effect of delayed-JOLs and sentence generation on children's monitoring accuracy and regulation of idiom study. *Metacognition and Learning*, 8(2), 173–191. doi:10.1007/s11409-013-9100-0
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker (Eds.), *Metacognition in educational theory and practice* (pp. 227–304). Mahwah, NJ: Erlbaum.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview.

*Educational Psychologist*, 25(1), 3-17. doi:10.1207/s15326985ep2501\_2

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background,

methodological developments, and future prospects. *American Educational Research Journal*,

45(1), 166–183. doi:10.3102/0002831207312909